THE PROBABILITY THAT A RANDOM MULTIGRAPH IS SIMPLE

SVANTE JANSON

Abstract. Consider a random multigraph G^* with given vertex degrees $d_1, \dots, d_n$, contructed by the configuration model. We show that, asymptotically for a sequence of such multigraphs with the number of edges $\frac{1}{2}\sum_i d_i \to \infty$, the probability that the multigraph is simple stays away from 0 if and only if $\sum_i d_i^2 = O(\sum_i d_i)$. This was previously known only under extra assumtions on the maximum degree $\max_i d_i$. We also give an asymptotic formula for this probability, extending previous results by several authors.

1. Introduction

If $n \geq 1$ and $(d_i)_1^n$ is a sequence of non-negative integers, we let $G(n, (d_i)_1^n)$ be the random (simple) graph with the n vertices $1, \dots, n$, and with vertex degrees $d_1, \dots, d_n$, uniformly chosen among all such graphs (provided that there are any such graphs at all; in particular, $\sum_i d_i$ has to be even). A standard method to study $G(n, (d_i)_1^n)$ is to consider the related random multigraph $G^*(n, (d_i)_1^n)$ defined by taking a set of d_i *half-edges* at each vertex i and then joining the half-edges into edges by taking a random partition of the set of all half-edges into pairs; see Section 2 for details. This is known as the configuration model, and such a partition of the half-edges is known as a *configuration*; this was introduced by Bollobás [2], see also Section II.4 of [3]. (See Bender and Canfield [1] and Wormald [14, 15] for related arguments.)

Note that $G^*(n, (d_i)_1^n)$ is defined for all $n \geq 1$ and all sequences $(d_i)_1^n$ such that $\sum_i d_i$ is even (we tacitly assume this throughout the paper), and that we obtain $G(n, (d_i)_1^n)$ if we condition $G^*(n, (d_i)_1^n)$ on being a simple graph. The idea of using the configuration method to study $G(n, (d_i)_1^n)$ is that $G^*(n, (d_i)_1^n)$ in many respects is a simpler object than $G(n, (d_i)_1^n)$; thus it is often possible to show results for $G(n, (d_i)_1^n)$ by first studying $G^*(n, (d_i)_1^n)$ and then conditioning on this multigraph being simple. It is then of crucial importance to be able to estimate the probability that $G^*(n, (d_i)_1^n)$ is simple, and in particular to decide whether

$$\liminf_{n\to\infty} \mathbb{P}\big(G^*(n, (d_i)_1^n) \text{ is simple}\big) > 0 \tag{1.1}$$

Date: September 28, 2006.

2000 *Mathematics Subject Classification.* 05C80; 05C30, 60C05.

for given sequences $(d_i)_1^n = (d_i^{(n)})_1^n$ (depending on $n \geq 1$). (Note that (1.1) implies that any statement holding for $G^*(n,(d_i)_1^n)$ with probability tending to 1 does so for $G(n,(d_i)_1^n)$ too.)

A natural condition that has been used by several authors using the configuration method (including myself [7]) as a sufficient condition for (1.1) is

$$\sum_{i=1}^n d_i = \Theta(n) \qquad \text{and} \qquad \sum_{i=1}^n d_i^2 = O(n) \tag{1.2}$$

together with some bound on $\max_i d_i$. (Recall that $A = \Theta(B)$ means that both $A = O(B)$ and $B = O(A)$ hold.) Results showing, or implying, that (1.2) and a condition on $\max_i d_i$ imply (1.1) have also been given by several authors, for example Bender and Canfield [1] with $\max_i d_i = O(1)$; Bollobás [2], see also Section II.4 in [3], with $\max_i d_i \leq \sqrt{2\log n} - 1$; McKay [10] with $\max_i d_i = o(n^{1/4})$; McKay and Wormald [13] with $\max_i d_i = o(n^{1/3})$. (Similar results have also been proved for bipartite graphs [9], digraphs [5], and hypergraphs [4].)

Indeed, it is not difficult to see that the method used by Bollobás [2, 3] works, assuming (1.2), provided only $\max_i d_i = o(n^{1/2})$, see Section 7. This has undoubtedly been noted by several experts, but we have not been able to find a reference to it in print when we have needed one.

One of our main result is that, in fact, (1.2) is sufficient for (1.1) without any assumption on $\max_i d_i$, even in cases where the Poisson approximation fails. Moreover, (1.2) is essentially necessary.

We remark that several papers (including several of the references given above) study $\mathbb{P}(G^*(n,(d_i)_1^n)$ is simple) from another point of view, namely by studying the number of simple graphs with given degree sequence $(d_i)_1^n$. It is easy to count configurations, and it follows that this number equals, with N the number of edges, see (1.3) below,

$$\frac{(2N)!}{2^N N! \prod_i d_i!} \mathbb{P}\bigl(G^*(n,(d_i)_1^n) \text{ is simple}\bigr);$$

such results are thus equivalent to results for $\mathbb{P}\bigl(G^*(n,(d_i)_1^n) \text{ is simple}\bigr)$. However, in this setting it is also interesting to obtain detailed asymptotics when $\mathbb{P}\bigl(G^*(n,(d_i)_1^n) \text{ is simple}\bigr) \to 0$; such results are included in several of the references above, but will not be treated here.

We will throughout the paper let N be the number of edges in $G^*(n,(d_i)_1^n)$. Thus

$$2N = \sum_{i=1}^n d_i. \tag{1.3}$$

It turns out that it is more natural to state our results in terms of N than n (the number of vertices). We can state our first result as follows; we use an index ν to emphasize that the result is asymptotic, and thus should be stated for a sequence (or another family) of multigraphs.

Theorem 1.1. *Consider a sequence of random multigraphs* $G^*_\nu = G^*\bigl(n_\nu, (d_i^{(\nu)})_1^n\bigr)$. *Let* $N_\nu = \frac{1}{2}\sum_i d_i^{(\nu)}$, *the number of edges in* G^*_ν, *and assume that, as* $\nu \to \infty$, $N_\nu \to \infty$. *Then*

(i) $\liminf_{\nu\to\infty} \mathbb{P}(G^*_\nu \text{ is simple}) > 0$ *if and only if* $\sum_i (d_i^{(\nu)})^2 = O(N_\nu)$;
(ii) $\lim_{\nu\to\infty} \mathbb{P}(G^*_\nu \text{ is simple}) = 0$ *if and only if* $\sum_i (d_i^{(\nu)})^2/N_\nu \to \infty$.

In the sequel we will for simplicity omit the index ν, but all results should be interpreted in the same way as Theorem 1.1.

Usually, one studies $G^*(n, (d_i)_1^n)$ as indexed by n. We then have the following special case of Theorem 1.1, which includes the claim above that (1.2) is sufficient for (1.1).

Corollary 1.2. *Let* $(d_i)_1^n = (d_i^{(n)})_1^n$ *be given for* $n \geq 1$. *Assume that* $N = \Theta(n)$. *Then, as* $n \to \infty$,

(i) $\liminf_{n\to\infty} \mathbb{P}\bigl(G^*(n, (d_i)_1^n) \text{ is simple}\bigr) > 0$ *if and only if* $\sum_i (d_i^{(\nu)})^2 = O(n)$,
(ii) $\mathbb{P}\bigl(G^*(n, (d_i)_1^n) \text{ is simple}\bigr) \to 0$ *if and only if* $\sum_i (d_i^{(\nu)})^2/n \to \infty$.

Remark 1.3. Although we have stated Corollary 1.2 as a special case of Theorem 1.1 with $N = O(n)$, it is essentially equivalent to Theorem 1.1. In fact, we may ignore all vertices of degree 0; thus we may assume that $d_i \geq 1$ for all i, and hence $2N \geq n$. If further $\sum_i d_i^2 = O(N)$, the Cauchy–Schwarz inequality yields

$$2N = \sum_{i=1}^n d_i \leq \Bigl(n \sum_{i=1}^n d_i^2\Bigr)^{1/2} = O(\sqrt{nN}),$$

and thus $N = \Theta(n)$. In the case $\sum_i d_i^2/n \to \infty$, it is possible to reduce some d_i to 1 such that then $N = \frac{1}{2}\sum_i d_i = \Theta(n)$ and still $\sum_i d_i^2/N \to \infty$; we omit the details since our proof does not use this route.

Our second main result is an asymptotic formula for the probability that $G^*(n, (d_i)_1^n)$ is simple.

Theorem 1.4. *Consider* $G^*(n, (d_i)_1^n)$ *and assume that* $N := \frac{1}{2}\sum_i d_i \to \infty$. *Let* $\lambda_{ij} := \sqrt{d_i(d_i-1)d_j(d_j-1)}/(2N)$; *in particular* $\lambda_{ii} = d_i(d_i-1)/(2N)$. *Then*

$$\mathbb{P}\bigl(G^*(n, (d_i)_1^n) \text{ is simple}\bigr) = \exp\Bigl(-\tfrac{1}{2}\sum_i \lambda_{ii} - \sum_{i<j}\bigl(\lambda_{ij} - \log(1+\lambda_{ij})\bigr)\Bigr) + o(1); \quad (1.4)$$

equivalently,

$$\mathbb{P}\big(G^*(n,(d_i)_1^n) \text{ is simple}\big)$$
$$= \exp\Big(-\frac{1}{4}\Big(\frac{\sum_i d_i^2}{2N}\Big)^2 + \frac{1}{4} + \frac{\sum_i d_i^2(d_i-1)^2}{16N^2} + \sum_{i<j}\big(\log(1+\lambda_{ij}) - \lambda_{ij} + \tfrac{1}{2}\lambda_{ij}^2\big)\Big) + o(1). \quad (1.5)$$

In many cases, $\sum_i d_i^2(d_i-1)^2$ in (1.5) may be replaced by the simpler $\sum_i d_i^4$; for example, this can be done whenever (1.1) holds, by Theorem 1.1 and (2.4). Note, however, that this is not always possible; a trivial counter example is obtained with $n = 1$ and $d_1 = 2N \to \infty$.

In the case $\max_i d_i = o(N^{1/2})$, Theorem 1.4 simplifies as follows; see also Section 7.

Corollary 1.5. *Assume that $N \to \infty$ and $\max_i d_i = o(N^{1/2})$. Let*

$$\Lambda := \frac{1}{2N}\sum_{i=1}^n \binom{d_i}{2} = \frac{\sum_i d_i^2}{4N} - \frac{1}{2}. \qquad (1.6)$$

Then

$$\mathbb{P}\big(G^*(n,(d_i)_1^n) \text{ is simple}\big) = \exp\big(-\Lambda - \Lambda^2\big) + o(1)$$
$$= \exp\Big(-\frac{1}{4}\Big(\frac{\sum_i d_i^2}{2N}\Big)^2 + \frac{1}{4}\Big) + o(1).$$

This formula is well known, at least under stronger conditions on $\max_i d_i$, see, for example, Bender and Canfield [1], Bollobás [3, Theorem II.16], McKay [10] and McKay and Wormald [13, Lemma 5.1].

2. Preliminaries

We introduce some more notation.

We will often write G^* for the random multigraph $G^*(n,(d_i)_1^n)$.

Let $V_n = \{1,\dots,n\}$; this is the vertex set of $G^*(n,(d_i)_1^n)$. We will in the sequel denote elements of V_n by u, v, w, possibly with indices. V_n is also the vertex set of the complete graph K_n, and we let E_n denote the edge set of K_n; thus E_n consists of the $\binom{n}{2}$ unordered pairs $\{v,w\}$, with $v, w \in V_n$ and $v \neq w$. We will use the notation vw for the edge $\{v,w\} \in E_n$.

For any multigraph G with vertex set V_n, and $u \in V_n$, we let $X_u(G)$ be the number of loops at u. Similarly, if $e = vw \in E_n$, we let $X_e(G) = X_{vw}(G)$ be the number of edges between v and w. We define further the indicators

$$I_u(G) := \mathbf{1}[X_u(G) \geq 1], \qquad u \in V_n,$$
$$J_e(G) := \mathbf{1}[X_e(G) \geq 2], \qquad e \in E_n,$$

and their sum

$$Y(G) := \sum_{u\in V_n} I_u(G) + \sum_{e\in E_n} J_e(G). \qquad (2.1)$$

Thus G is a simple graph if and only if $Y(G) = 0$, and our task is to estimate $\mathbb{P}(Y(G^*) = 0)$.

As said above, the idea of the configuration model is that we fix a set of d_v half-edges for every vertex v; we denote these half-edges by $v^{(1)}, \dots, v^{(d_v)}$, and say that they belong to v, or are at v. These sets are assumed to be disjoint, so the total number of half-edges is $\sum_v d_v = 2N$. A configuration is a partition of the $2N$ half-edges into N pairs, and each configuration defines a multigraph with vertex set V_n and vertex degrees d_v by letting every pair $\{x, y\}$ of half-edges in the configuration define an edge; if x is a half-edge at v and y is a half-edge at w, we form an edge between v and w (and thus a loop if $v = w$). We express this construction by saying that we join the two half-edges x and y to an edge; we may denote this edge by xy. Recall that G^* is the random multigraph obtain from a (uniform) random configuration by this construction.

We will until Section 6 assume that

$$\sum_v d_v^2 = O(N), \tag{2.2}$$

i.e., that $\sum_v d_v^2 \le CN$ for some constant C. (The constants implicit in the estimates below may depend on this constant C.) Note that an immediate consequence is

$$\max_v d_v = O(N^{1/2}) = o(N). \tag{2.3}$$

We may thus assume that N is so large that $\max_v d_v < N/10$, say, and thus all terms like $N - d_v$ are of order N. (The estimates we will prove are trivially true for any finite number of N by taking the implicit constants large enough; thus it suffices to prove them for large N.)

Note further that (2.2) implies, using (2.3), that for any fixed $k \ge 2$

$$\sum_v d_v^k \le (\max_v d_v)^{k-2} \sum_v d_v^2 = O(N^{k/2}). \tag{2.4}$$

We further note that we can assume $d_v \ge 1$ for all v, since vertices with degree 0 may be removed without any difference to our results. (This is really not necessary, but it means that we do not even have to think about, for example, d_v^{-1} in some formulas below.)

We will repeatedly use the *subsubsequence principle*, which says that if $(x_n)_n$ is a sequence of real numbers and a is a number such that every subsequence of $(x_n)_n$ has a subsequence that converges to a, then the full sequence converges to a. (This holds in any topological space.)

We denote the falling factorials by $x^{\underline{k}} := x(x-1)\cdots(x-k+1)$.

3. Two probabilistic lemmas

We will use two simple probabilistic lemmas. The first is (at least part (i)) a standard extension of the inclusion-exclusion principle; we include a proof for completeness.

Lemma 3.1. *Let W be a non-negative integer-valued random variable such that $\mathbb{E}\, R^W < \infty$ for some $R > 2$.*

(i) *Then, for every $j \geq 0$,*

$$\mathbb{P}(W = j) = \sum_{k=j}^{\infty} (-1)^{k-j} \binom{k}{j} \frac{1}{k!}\, \mathbb{E}(W^{\underline{k}}).$$

(ii) *More generally, for every random variable Z such that $\mathbb{E}(|Z| R^W) < \infty$ for some $R > 2$, and every $j \geq 0$,*

$$\mathbb{E}(Z \cdot \mathbf{1}[W = j]) = \sum_{k=j}^{\infty} (-1)^{k-j} \binom{k}{j} \frac{1}{k!}\, \mathbb{E}(Z W^{\underline{k}}).$$

Proof. For (i), let $f(t) := \mathbb{E}(t^W) = \sum_j \mathbb{P}(W = j) t^j$ be the probability generating function of W; this is by assumption convergent for $|t| \leq R$, at least. If $|t| \leq R - 1$ we have

$$f(t+1) = \mathbb{E}(1+t)^W = \sum_{k=0}^{\infty} \mathbb{E} \binom{W}{k} t^k$$

and thus, if $|t| \leq R - 2$,

$$f(t) = \sum_{k=0}^{\infty} \mathbb{E} \binom{W}{k} (t-1)^k = \sum_{k=0}^{\infty} \mathbb{E}\big(W^{\underline{k}}/k!\big) \sum_{j=0}^{\infty} \binom{k}{j} t^j (-1)^{j-k}. \tag{3.1}$$

The double series is absolutely convergent since

$$\sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \mathbb{E}\big(W^{\underline{k}}/k!\big) \binom{k}{j} |t|^j = \sum_{k=0}^{\infty} \mathbb{E} \binom{W}{k} (|t|+1)^k = f(|t|+2) < \infty.$$

Hence the result follows by extracting the coefficients of t^j in (3.1)

Part (ii) is proved it the same way, using instead $f(t) := \mathbb{E}(Z t^W)$. □

The next lemma could be proved by Lemma 3.1 if made the hypothesis somewhat stronger, but we prefer another proof.

Lemma 3.2. *Let $(W_\nu)_\nu$ and $(\widetilde{W}_\nu)_\nu$ be two sequences of non-negative integer-valued random variables such that, for some $R > 1$*

$$\sup_\nu \mathbb{E}\big(R^{W_\nu}\big) < \infty \tag{3.2}$$

and, for each fixed $k \geq 1$,

$$\mathbb{E}\big(W_\nu^{\underline{k}}\big) - \mathbb{E}\big(\widetilde{W}_\nu^{\underline{k}}\big) \to 0 \qquad \text{as } \nu \to \infty. \tag{3.3}$$

Then, as $\nu \to \infty$,

$$\mathbb{P}(W_\nu = 0) - \mathbb{P}(\widetilde{W}_\nu = 0) \to 0. \tag{3.4}$$

Proof. By the subsubsequence principle, it suffices to prove that every subsequence has a subsequence along which (3.4) holds. Since (3.2) implies that the sequence $(W_\nu)_\nu$ is tight, we can by selecting a suitable subsequence assume that $W_\nu \xrightarrow{\mathrm{d}} W$ for some random variable W (see Sections 5.8.2 and 5.8.3 in Gut [6]). Moreover, (3.2) implies uniform integrability of the powers W_ν^k for each k, and we thus have, as $\nu \to \infty$ along the selected subsequence, $\mathbb{E}(W_\nu^k) \to \mathbb{E}(W^k)$ for every k and thus also $\mathbb{E}\big(W_\nu^{\underline{k}}\big) \to \mathbb{E}\big(W^{\underline{k}}\big)$ (see Theorems 5.4.2 and 5.5.9 in [6]). By (3.3), this yields also $\mathbb{E}\big(\widetilde{W}_\nu^{\underline{k}}\big) \to \mathbb{E}\big(W^{\underline{k}}\big)$. Furthermore, (3.2) implies by Fatou's lemma (Theorem 5.5.8 in [6]) that $\mathbb{E}(R^W) \le \liminf \mathbb{E}(R^{W_\nu}) < \infty$, or $\mathbb{E}\big(e^{tW}\big) < \infty$ with $t = \log R > 0$; hence the distribution of W is determined by its moments (see Section 4.10 in [6]). Consequently, by the method of moments (Theorem 6.7 in [8]), still along the subsequence, $\widetilde{W}_\nu \xrightarrow{\mathrm{d}} W$ and thus

$$\mathbb{P}(W_\nu = 0) - \mathbb{P}(\widetilde{W}_\nu = 0) \to \mathbb{P}(W = 0) - \mathbb{P}(W = 0) = 0. \qquad \square$$

Remark 3.3. The same proof gives the stronger statement

$$d_{\mathrm{TV}}(W_\nu, \widetilde{W}_\nu) := \sum_j |\,\mathbb{P}(W_\nu = j) - \mathbb{P}(\widetilde{W}_\nu = j)| \to 0.$$

4. Individual probabilities

We begin by estimating the probabilities $\mathbb{P}(I_u(G^*) = 1)$ and $\mathbb{P}(J_{vw}(G^*) = 1)$. The following form will be convenient.

Lemma 4.1. *Suppose $\sum_v d_v^2 = O(N)$. Then, for G^*, and for all $u, v, w \in V_n$, if N is so large that $d_u \le N$,*

$$-\log \mathbb{P}(I_u = 0) = -\log \mathbb{P}(X_u = 0) = \frac{d_u(d_u - 1)}{4N} + O\left(\frac{d_u^3}{N^2}\right)$$

and, with $\lambda_{vw} := \sqrt{d_v(d_v - 1)d_w(d_w - 1)}/(2N)$ as in Theorem 1.4,

$$\begin{aligned} -\log \mathbb{P}(J_{vw} = 0) &= -\log \mathbb{P}(X_{vw} \le 1) \\ &= -\log(1 + \lambda_{vw}) + \lambda_{vw} + O\left(\frac{(d_v + d_w)d_v^2 d_w^2}{N^3}\right). \end{aligned}$$

Proof. The calculation for loops is simple. We construct the random configuration by first choosing partners to the half-edges at u, one by one. A simple counting shows that

$$\mathbb{P}(X_u = 0) = \prod_{i=1}^{d_u} \left(1 - \frac{d_u - i}{2N - 2i + 1}\right)$$

and thus, for large N, using $-\log(1 - x) = x + O(x^2)$ when $|x| \le 1/2$,

$$-\ln \mathbb{P}(X_u = 0) = \sum_{i=1}^{d_u} \left(\frac{d_u - i}{2N} + O\left(\frac{d_u^2}{N^2}\right)\right) = \frac{d_u(d_u - 1)}{4N} + O\left(\frac{d_u^3}{N^2}\right).$$

For multiple edges, a similar direct approach would be much more complicated because of the possibility of loops at v or w. We instead use Lemma 3.1(i), with $W = X_{vw}$. We may assume $d_v, d_w \geq 2$, since the result otherwise is trivial. $X_{vw}^{\underline{k}}$ is the number of ordered k-tuples of edges between v and w; the corresponding pairs in the configuration may be chosen in $d_v^{\underline{k}} d_w^{\underline{k}}$ ways, and each such set of k pairs appears in the configuration with probability $((2N-1)(2N-3)\cdots(2N-2k+1))^{-1}$. Thus

$$\mathbb{E}\, X_{vw}^{\underline{k}} = \frac{d_v^{\underline{k}} d_w^{\underline{k}}}{(2N-1)\cdots(2N-2k+1)} = \frac{d_v^{\underline{k}} d_w^{\underline{k}}}{2^k (N-1/2)^{\underline{k}}} = \prod_{i=0}^{k-1} \frac{(d_v-i)(d_w-i)}{2(N-1/2-i)}. \tag{4.1}$$

In particular,

$$\mathbb{E}\, X_{vw}^{\underline{2}} = \lambda_{vw}^2 (1 + O(1/N)) \tag{4.2}$$

and if $k \geq 3$, uniformly in k,

$$\mathbb{E}\, X_{vw}^{\underline{k}} = (1 + O(1/N)) \lambda_{vw}^2 \prod_{i=2}^{k-1} \frac{(d_v-i)(d_w-i)}{2(N-1/2-i)}. \tag{4.3}$$

Since $d_v < N$ (for large N at least, see (2.3)), the ratios $(d_v - i)/(N - 1/2 - i)$ decrease as i increases; hence, for large N and $i \geq 2$,

$$\frac{d_v - i}{2(N-i-1/2)} \leq \frac{d_v - 2}{2N-5} < \frac{d_v - 1}{2N} < \frac{\sqrt{d_v(d_v-1)}}{2N}$$

and (4.3) yields, uniformly for $k \geq 3$,

$$\mathbb{E}\, X_{vw}^{\underline{k}} \leq (1 + O(1/N)) \lambda_{vw}^2 \prod_{i=2}^{k-1} \lambda_{vw} = (1 + O(1/N)) \lambda_{vw}^k. \tag{4.4}$$

In the opposite direction, still uniformly for $k \geq 3$,

$$\begin{aligned}\mathbb{E}\, X_{vw}^{\underline{k}} / \lambda_{vw}^k &\geq \prod_{i=2}^{k-1} \frac{(d_v-i)(d_w-i)}{2N\lambda_{vw}} \geq \prod_{i=2}^{k-1} \frac{(d_v-i)(d_w-i)}{d_v d_w} \\ &= \prod_{i=2}^{k-1} (1 - i/d_v)(1 - i/d_w) \geq 1 - \sum_{i=2}^{k-1} (i/d_v + i/d_w) \geq 1 - \frac{k^2}{d_v} - \frac{k^2}{d_w}.\end{aligned}$$

Together with (4.4), this shows that, uniformly for $k \geq 3$,

$$\mathbb{E}\, X_{vw}^{\underline{k}} = \lambda_{vw}^k \left(1 + O\left(k^2 (d_v^{-1} + d_w^{-1})\right)\right). \tag{4.5}$$

We now use Lemma 3.1(i), noting that trivially $\mathbb{E}\, R^{X_{vw}} \leq R^{d_v} < \infty$ for every R. Thus, using (4.2) and (4.5) and observing that $\lambda_{vw} = O(1)$ by

(2.3),

$$\begin{aligned}\mathbb{P}(X_{vw}=0) &= 1-\mathbb{E}\,X_{vw}+\tfrac{1}{2}\lambda_{vw}^2(1+O(N^{-1}))\\ &\qquad+\sum_{k=3}^{\infty}(-1)^k\frac{\lambda_{vw}^k}{k!}\left(1+O\left(k^2(d_v^{-1}+d_w^{-1})\right)\right)\\ &= 1-\mathbb{E}\,X_{vw}+e^{-\lambda_{vw}}-1+\lambda_{vw}+O\left(\lambda_{vw}^2N^{-1}+\lambda_{vw}^3e^{\lambda_{vw}}(d_v^{-1}+d_w^{-1})\right)\\ &= e^{-\lambda_{vw}}+\lambda_{vw}-\mathbb{E}\,X_{vw}+O\left((d_v^2d_w^3+d_v^3d_w^2)N^{-3}\right).\end{aligned}$$

Similarly, still by Lemma 3.1(i),

$$\begin{aligned}\mathbb{P}(X_{vw}=1) &= \mathbb{E}\,X_{vw}-\lambda_{vw}^2(1+O(N^{-1}))\\ &\qquad+\sum_{k=3}^{\infty}(-1)^{k-1}\frac{\lambda_{vw}^k}{(k-1)!}\left(1+O\left(k^2(d_v^{-1}+d_w^{-1})\right)\right)\\ &= \mathbb{E}\,X_{vw}+\lambda_{vw}(e^{-\lambda_{vw}}-1)+O\left(\lambda_{vw}^2N^{-1}+\lambda_{vw}^3e^{\lambda_{vw}}(d_v^{-1}+d_w^{-1})\right)\\ &= \mathbb{E}\,X_{vw}+\lambda_{vw}e^{-\lambda_{vw}}-\lambda_{vw}+O\left((d_v^2d_w^3+d_v^3d_w^2)N^{-3}\right).\end{aligned}$$

Summing these two equations we find

$$\mathbb{P}(X_{vw}\le 1)=(1+\lambda_{vw})e^{-\lambda_{vw}}+O\left((d_v^2d_w^3+d_v^3d_w^2)N^{-3}\right)$$

and the result follows, noting that $(1+\lambda_{vw})e^{-\lambda_{vw}}$ is bounded below since $\lambda_{vw}=O(1)$. ∎

5. Joint probabilities

Our goal is to show that the indicators $I_u(G^*)$ and $J_e(G^*)$ are almost independent for different u and e; this is made precise in the following lemma.

We define for convenience, for $u\in V_n$ and $e=vw\in E_n$,

$$\mu_u := d_u^2/N \qquad\text{and}\qquad \mu_e := d_vd_w/N. \tag{5.1}$$

It follows easily from (4.1) and a similar calculation for loops that

$$\mathbb{E}\big(X_u(G^*)^{\underline{k}}\big)\le\mu_u^k \qquad\text{and}\qquad \mathbb{E}\big(X_e(G^*)^{\underline{k}}\big)\le\mu_e^k, \qquad k\ge 1. \tag{5.2}$$

In particular, omitting the argument G^*,

$$\begin{aligned}\mathbb{P}(I_u=1)&=\mathbb{E}\,I_u\le\mathbb{E}\,X_u\le\mu_u,\\ \mathbb{P}(J_e=1)&=\mathbb{E}\,J_e\le\mathbb{E}\,X_e^{\underline{2}}\le\mu_e^2.\end{aligned} \tag{5.3}$$

More precisely, it follows easily from Lemma 4.1 that (for N large at least) $\mathbb{P}(I_u=1)=\Theta(\mu_u)$ and $\mathbb{P}(J_e=1)=\Theta(\mu_e^2)$ provided $d_v,d_w\ge 2$; this may help understanding our estimates but will not be used below.

Lemma 5.1. *Suppose $\sum_v d_v^2=O(N)$. Let $l\ge 0$ and $m\ge 0$ be fixed. For any sequences of distinct vertices $u_1,\dots,u_l\in V_n$ and edges $e_1,\dots,e_m\in E_n$,*

let $e_j = v_j w_j$ *and let* F *be the set of vertices that appear at least twice in the list* $u_1, \dots, u_l, v_1, w_1, \dots, v_m, w_m$. *Then,*

$$\mathbb{E}\Bigl(\prod_{i=1}^{l} I_{u_i}(G^*) \prod_{j=1}^{m} J_{e_j}(G^*)\Bigr) = \prod_{i=1}^{l} \mathbb{E}(I_{u_i}(G^*)) \prod_{j=1}^{m} \mathbb{E}(J_{e_j}(G^*)) + O\Bigl(\Bigl(N^{-1} + \sum_{v \in F} d_v^{-1}\Bigr) \prod_{i=1}^{l} \mu_{u_i} \prod_{j=1}^{m} \mu_{e_j}^2\Bigr). \quad (5.4)$$

The implicit constant in the error term may depend on l and m but not on $(u_i)_i$ and $(e_j)_j$. All similar statements below are to be interpreted similarly.

The proof of Lemma 5.1 is long, and contains several other lemmas. The idea of the proof is to use induction in $l + m$. In the inductive step we select one of the indicators, J_{e_1} say, and then show that the product of the other indicators is almost independent of X_{e_1}, and thus of J_{e_1}. In order to do so, we would like to condition on the value of X_{e_1}. But the effects of conditioning on $X_{e_1} = k$ are complicated and we find it difficult to argue directly with these conditionings (see Remark 5.7). Therefore, we begin with another, related but technically much simpler conditioning.

Fix two distinct vertices v and w. For $0 \le k \le \min(d_v, d_w)$, let $\mathcal{E}_k$ be the event that the random configuration contains the k pairs of half-edges $\{v^{(i)}, w^{(i)}\}$, $1 \le i \le k$, and let the corresponding random multigraph, i.e., G^* conditioned on $\mathcal{E}_k$, be denoted G^*_k. G^*_k thus contains at least k edges between v and w, but there may be more. Note that $G^*_0 = G^*$.

We begin with an estimate related to Lemma 5.1, but cruder.

Lemma 5.2. *Suppose* $\sum_v d_v^2 = O(N)$. *Let* l, m *and* $r_1, \dots, r_l, s_1, \dots, s_m$ *be fixed non-negative integers. For any sequences of distinct vertices* $u_1, \dots, u_l \in V_n$ *and edges* $e_1, \dots, e_m \in E_n$,

$$\mathbb{E}\Bigl(\prod_{i=1}^{l} X_{u_i}(G^*)^{\underline{r_i}} \prod_{j=1}^{m} X_{e_j}(G^*)^{\underline{s_j}}\Bigr) = O\Bigl(\prod_{i=1}^{l} \mu_{u_i}^{r_i} \prod_{j=1}^{m} \mu_{e_j}^{s_j}\Bigr). \quad (5.5)$$

In particular,

$$\mathbb{E}\Bigl(\prod_{i=1}^{l} I_{u_i}(G^*) \prod_{j=1}^{m} J_{e_j}(G^*)\Bigr) = O\Bigl(\prod_{i=1}^{l} \mu_{u_i} \prod_{j=1}^{m} \mu_{e_j}^2\Bigr). \quad (5.6)$$

The estimates (5.5) *and* (5.6) *hold with* G^* *replaced by* G^*_k *too, uniformly in* k, *provided the edges* $e_1, \dots, e_m$ *are distinct from the edge* vw *used to define* G^*_k. *If* vw *equals some* e_j, *then* (5.5) *still holds for* G^*_k, *if we replace* X_{e_j} *by* $X_{e_j} - k$ *when* $e_j = vw$.

Proof. We argue as for (4.1). Let, again, $e_j = v_j w_j$ and let $t := r_1 + \dots + r_l + s_1 + \dots + s_m$. The expectation in (5.5) is the number of t-tuples of disjoint pairs of half-edges such that the first r_1 pairs have both half-edges

belonging to u_1, and so on, until the last s_m that each consist of one half-edge at v_m and one at w_m, times the probability that a given such t-tuple is contained in a random configuration. The number of such t-tuples is at most $\prod_{i=1}^{l} d_{u_i}^{2r_i} \prod_{j=1}^{m} (d_{v_j} d_{w_j})^{s_j}$ and the probability is $((2N-1)\cdots(2N-2t+1))^{-1} < N^{-t}$ (provided $N \geq 2t$). The estimate (5.5) follows, recalling (5.1), and (5.6) is an immediate consequence since $I_u \leq X_u$ and $J_e \leq X_e^{\underline{2}}$.

The same argument proves the estimates for G_k^*. There is a minor change in the probability above, replacing N by $N-2k$; nevertheless, the estimates are uniform in k because $k \leq d_v = O(N^{1/2})$. (There may also be some t-tuples that are excluded because they clash with the special pairs $\{v^{(i)}, w^{(i)}\}$, $i = 1, \dots, k$; this only helps.) ◻

Let $u_1, \dots, u_l \in V_n$ and $e_1, \dots, e_m \in E_n$ be as in Lemma 5.1, and assume that $m \geq 1$. We choose $v = v_1$ and $w = w_1$, so $e_1 = vw$, for the definition of G_k^*.

If $k \geq 1$, we can couple G_k^* and G_{k-1}^* as follows. Start with a random configuration containing the k special pairs $\{v^{(i)}, w^{(i)}\}$. Then select, at random, a half-edge x among all half-edges except $v^{(1)}, \dots, v^{(k)}, w^{(1)}, \dots, w^{(k-1)}$. If $x = w^{(k)}$ do nothing. Otherwise, let y be the half-edge paired to x; remove the two pairs $\{v^{(k)}, w^{(k)}\}$ and $\{x, y\}$ from the configuration and replace them by $\{v^{(k)}, x\}$ and $\{w^{(k)}, y\}$. (This is called a switching; see McKay [10] and McKay and Wormald [11, 13] for different but related arguments with switchings.)

It is clear that this gives a configuration in $\mathcal{E}_{k-1}$ with the correct uniform distribution. Passing to the multigraphs, we thus obtain a coupling of G_k^* and G_{k-1}^* such that the two multigraphs differ (if at all) in that one edge between v and w and one other edge have been deleted, and two new edges are added, one at v and one at w.

Let Z denote the product $\prod_{i=1}^{l} I_{u_i} \prod_{j=2}^{m} J_{e_j}$ of the chosen indicators except J_{e_1}. Define $F_1 \subseteq \{v, w\}$ to be the set of endpoints of $vw = e_1$ that also appear as some u_i or as an end-point of some other e_j; thus $F_1 = F \cap \{v, w\}$. We claim the following.

Lemma 5.3. *Suppose $\sum_v d_v^2 = O(N)$. With notations as above, uniformly in k with $1 \leq k \leq \min(d_v, d_w)$,*

$$\mathbb{E}\big(Z(G_k^*)\big) - \mathbb{E}\big(Z(G_{k-1}^*)\big) = O\bigg(\Big(N^{-1} + \sum_{v \in F_1} d_v^{-1}\Big) \prod_{i=1}^{l} \mu_{u_i} \prod_{j=2}^{m} \mu_{e_j}^2\bigg). \quad (5.7)$$

Proof. We use the coupling above. Recall that $Z = 0$ or 1, so if $Z(G_k^*)$ and $Z(G_{k-1}^*)$ differ, then one of them equals 0 and the other equals 1.

First, if $Z(G_k^*) = 1$ and $Z(G_{k-1}^*) = 0$, then the edge xy deleted from G_k^* must be either the only loop at some u_i, or one of exactly two edges between v_j and w_j for some $j \geq 2$. Hence, for any configuration with $Z(G_k^*) = 1$, there are less than $l + 2m$ such edges, and the probability that one of them

is destroyed is less than $(l+2m)/(N-k)=O(1/N)$. Hence,

$$\mathbb{P}\big(Z(G_k^*)>Z(G_{k-1}^*)\big)=O\big(\mathbb{E}(Z(G_k^*))/N\big). \tag{5.8}$$

Define $M:=\prod_{i=1}^{l}\mu_{u_i}\prod_{j=2}^{m}\mu_{e_j}^2$. By Lemma 5.2, $\mathbb{E}\,Z(G_k^*)=O(M)$, so the probability in (5.8) is $O(M/N)$, which is dominated by the right-hand side of (5.7).

In the opposite direction, $Z(G_k^*)=0$ and $Z(G_{k-1}^*)=1$ may happen in several ways. We list the possibilities as follows. (It is necessary but not necessarily sufficient for $Z(G_k^*)<Z(G_{k-1}^*)$ that one of them holds.)

(i) v is an endpoint of one of the edges $e_2,\dots,e_m$, say $v=v_2$ so $e_2=vw_2$; the new edge from v goes to w_2; there already is (exactly) one edge between v and w_2 in G_k^*; if we write $Z'=\prod_{i=1}^{l}I_{u_i}\prod_{j=3}^{m}J_{e_j}$, so that $Z=J_{e_2}Z'$, then $Z'(G_k^*)=1$.
(ii) v equals one of $u_1,\dots,u_l$, say $v=u_1$; the new edge from v is a loop; if we write $Z'=\prod_{i=2}^{l}I_{u_i}\prod_{j=2}^{m}J_{e_j}$, so that $Z=I_{u_1}Z'$, then $Z'(G_k^*)=1$.
(iii) Two similar cases with v replaced by w.
(iv) Both v and w are endpoints of edges e_j, say $v=v_2$ and $w=w_3$, so that $e_2=vw_2$ and $e_3=wv_3$; the two new edges go from v to w_2 and from w to v_3; there are already such edges in G_k^*; if $Z''=\prod_{i=1}^{l}I_{u_i}\prod_{j=4}^{m}J_{e_j}$, so that $Z=J_{e_2}J_{e_3}Z''$, then $Z''(\overline{G}_k^*)=1$, where $\overline{G}_k^*$ is G_k^* with one edge between w_2 and v_3 deleted.
(v) Both v and w equal some u_i, say $v=u_1$ and $w=u_2$; the new edges from v and w are loops; if $Z''=\prod_{i=3}^{l}I_{u_i}\prod_{j=2}^{m}J_{e_j}$, so that $Z=I_{u_1}I_{u_2}Z''$, then $Z''(G_k^*)=1$.
(vi) A similar mixed case where, say $v=v_2$ and $w=u_1$.
(vii) The same with v and w interchanged.

Consider case (i). For any configuration, the probability that the new edge from v goes to w_2 is $d_{w_2}/(2N-2k+1)=O(d_{w_2}/N)$. Since we also need $Z'(G_k^*)=1$ and $X_{e_2}(G_k^*)\ge 1$, the probability of case (i) is at most

$$O\big(d_{w_2}N^{-1}\,\mathbb{E}\big(X_{e_2}(G_k^*)Z'(G_k^*)\big)\big).$$

Now, by Lemma 5.2, for convenience omitting the arguments G_k^* here and often below in this proof,

$$\mathbb{E}\big(X_{e_2}Z'\big)\le\mathbb{E}\bigg(\prod_{i=1}^{l}X_{u_i}X_{e_2}\prod_{j=3}^{m}X_{e_j}^{\underline{2}}\bigg)=O\bigg(\prod_{i=1}^{l}\mu_{u_i}\mu_{e_2}\prod_{j=3}^{m}\mu_{e_j}^2\bigg)=O\big(M/\mu_{e_2}\big).$$

Moreover, $d_{w_2}/N=\mu_{e_2}/d_v$, so the probability of case (i) is $O(M/d_v)$; note that the case only can happen if $v\in F_1$, so this is covered by the right-hand side of (5.7).

Case (ii) is similar (and slightly simpler).

Case (iii) occurs, by symmetry, with probability $O(M/d_w)$, and only if $w\in F_1$.

In case (iv), the other destroyed edge must go between w_2 and v_3. For any configuration, the probability that such an edge is chosen is $O(X_{w_2v_3}/N)$. We study two subcases. If one of the edges $e_4,\dots,e_m$ equals w_2v_3, say $e_4 = w_2v_3$, then we, moreover, need at least three edges between w_2 and v_3 in G^*_k, since one of them is destroyed. We also need $X_{e_2}(G^*_k) \ge 1$ and $X_{e_3}(G^*_k) \ge 1$. Thus the probability of this case then is

$$O\Big(N^{-1}\,\mathbb{E}\big(X_{w_2v_3}X_{e_2}X_{e_3}Z''(\overline{G}^*_k)\big)\Big) = O\Big(N^{-1}\,\mathbb{E}\big(X_{e_2}X_{e_3}X_{e_4}\mathbf{1}[X_{e_4}\ge 3]Z''\big)\Big).$$

By Lemma 5.2 we have

$$\begin{aligned}\mathbb{E}\big(X_{e_2}X_{e_3}X_{e_4}\mathbf{1}[X_{e_4}\ge 3]Z''\big) &\le \mathbb{E}\bigg(\prod_{i=1}^{l} X_{u_i}\cdot X_{e_2}X_{e_3}X_{e_4}^{\underline{3}}\prod_{j=5}^{m} X_{e_j}^{\underline{2}}\bigg)\\ &= O\bigg(\prod_{i=1}^{l}\mu_{u_i}\cdot\mu_{e_2}\mu_{e_3}\mu_{e_4}^3\prod_{j=5}^{m}\mu_{e_j}^2\bigg)\\ &= O\big(M\mu_{e_4}/(\mu_{e_2}\mu_{e_3})\big).\end{aligned}$$

In this case we have $\mu_{e_2}\mu_{e_3} = \mu_{e_1}\mu_{e_4}$, so the probability is $O(N^{-1}M/\mu_{e_1}) = O(M/d_vd_w)$.

In the second subcase, w_2v_3 does not equal any of $e_4,\dots,e_m$. We then obtain similarly the probability

$$\begin{aligned}O\Big(N^{-1}\,\mathbb{E}\big(X_{w_2v_3}X_{e_2}X_{e_3}Z''\big)\Big) &= O\bigg(N^{-1}\prod_{i=1}^{l}\mu_{u_i}\cdot\mu_{w_2v_3}\mu_{e_2}\mu_{e_3}\prod_{j=4}^{m}\mu_{e_j}^2\bigg)\\ &= O\big(N^{-1}M\mu_{w_2v_3}/(\mu_{e_2}\mu_{e_3})\big),\end{aligned}$$

which again equals $O(M/(d_vd_w))$. Finally, note that in case (iv), $F_1 = \{v,w\}$.

In case (v), the other destroyed edge is also an edge between v and w; given a configuration, the probability of this is $O((X_{vw}-k)/N)$. The probability of case (v) is thus

$$\begin{aligned}O\Big(N^{-1}\,\mathbb{E}\big((X_{vw}-k)Z''\big)\Big) &= O\bigg(N^{-1}\mu_{vw}\prod_{i=3}^{l}\mu_{u_i}\prod_{j=1}^{m}\mu_{e_j}^2\bigg)\\ &= O\big(M/(d_vd_w)\big).\end{aligned}$$

$F_1 = \{v,w\}$ in case (v) too.

Cases (vi) and (vii) are similar to case (iv), and lead to the same estimate. Again $F_1 = \{v,w\}$.

By (5.8) and our estimates for the different cases above, the probability that $Z(G^*_k)$ and $Z(G^*_{k-1})$ differ is bounded by the right-hand side of (5.7), which completes the proof. $\square$

We can now estimate the expectation of $Z(G^*_k)$ conditioned on the value of $X_{vw}(G^*_k)$. We state only the result we need. (See also (5.12). These results can be rewritten as estimates of conditional expectations.)

Lemma 5.4. *Suppose $\sum_v d_v^2 = O(N)$. With notations as above,*

$$\begin{aligned}&\mathbb{E}\bigl(Z(G^*)J_{e_1}(G^*)\bigr)\\&\quad= \mathbb{E}\bigl(Z(G^*)\bigr)\,\mathbb{E}\bigl(J_{e_1}(G^*)\bigr) + O\biggl(\Bigl(N^{-1} + \sum_{v\in F_1} d_v^{-1}\Bigr)\prod_{i=1}^{l}\mu_{u_i}\prod_{j=1}^{m}\mu_{e_j}^2\biggr).\end{aligned} \tag{5.9}$$

Proof. We can write $X_{vw}(G^*)^{\underline{k}} = \sum_{\alpha\in\mathcal{A}} I_\alpha$, where $\mathcal{A}$ is the set of all ordered k-tuples of disjoint pairs (x,y) of half-edges with x belonging to v and y to w, and I_α is the indicator that the k pairs in α all belong to the configuration. By symmetry, $\mathbb{E}\bigl(Z(G^*)\mid I_\alpha=1\bigr)$ is the same for all $\alpha\in\mathcal{A}$; since G^*_k is obtained by conditioning G^* on I_α for a specific α, we thus have $\mathbb{E}\bigl(Z(G^*)\mid I_\alpha=1\bigr) = \mathbb{E}\bigl(Z(G^*_k)\bigr)$ for all α. Consequently,

$$\begin{aligned}\mathbb{E}\Bigl(X_{vw}(G^*)^{\underline{k}}Z(G^*)\Bigr) &= \sum_{\alpha\in\mathcal{A}}\mathbb{E}\bigl(I_\alpha Z(G^*)\bigr) = \sum_{\alpha\in\mathcal{A}}\mathbb{E}\bigl(Z(G^*)\mid I_\alpha=1\bigr)\,\mathbb{P}(I_\alpha=1)\\&= \mathbb{E}\bigl(Z(G^*_k)\bigr)\sum_{\alpha\in\mathcal{A}}\mathbb{E}\,I_\alpha = \mathbb{E}\bigl(Z(G^*_k)\bigr)\,\mathbb{E}\bigl(X_{vw}(G^*)^{\underline{k}}\bigr).\end{aligned} \tag{5.10}$$

We write the error term on the right-hand side of (5.7) as $O(R)$. Since (5.7) is uniform in k, and $G^*_0 = G^*$, Lemma 5.3 yields

$$\mathbb{E}\bigl(Z(G^*_k)\bigr) = \mathbb{E}\bigl(Z(G^*)\bigr) + O(kR). \tag{5.11}$$

We now use Lemma 3.1(ii) and (i) and find, for any j, using (5.10) and (5.11),

$$\begin{aligned}\mathbb{E}\bigl(Z(G^*)\cdot\mathbf{1}[X_{vw}(G^*)=j]\bigr) &= \sum_{k=j}^{\infty}(-1)^{k-j}\binom{k}{j}\frac{1}{k!}\,\mathbb{E}\bigl(X_{vw}(G^*)^{\underline{k}}Z(G^*)\bigr)\\&= \sum_{k=j}^{\infty}(-1)^{k-j}\binom{k}{j}\frac{1}{k!}\,\mathbb{E}\bigl(X_{vw}(G^*)^{\underline{k}}\bigr)\bigl(\mathbb{E}\,Z(G^*)+O(kR)\bigr)\\&= \mathbb{E}\bigl(Z(G^*)\bigr)\,\mathbb{P}\bigl(X_{vw}(G^*)=j\bigr) + O\left(\sum_{k=j}^{\infty}\binom{k}{j}\frac{1}{k!}\,\mathbb{E}\bigl(X_{vw}(G^*)^{\underline{k}}\bigr)kR\right).\end{aligned}$$

By (5.2), the sum inside the last O is at most

$$\sum_{k=j}^{\infty}\binom{k}{j}\frac{k}{k!}\mu_{vw}^kR = \sum_{l=0}^{\infty}\frac{j+l}{j!\,l!}\mu_{vw}^{j+l}R = \Bigl(\frac{\mu_{vw}^j}{(j-1)!}+\frac{\mu_{vw}^{j+1}}{j!}\Bigr)e^{\mu_{vw}}R.$$

Since $\mu_{vw} = O(1)$ by (2.3), we thus find, uniformly in $j \geq 1$,

$$\begin{aligned} &\mathbb{E}\bigl(Z(G^*) \cdot \mathbf{1}[X_{vw}(G^*) = j]\bigr) \\ &\qquad = \mathbb{E}\bigl(Z(G^*)\bigr)\, \mathbb{P}\bigl(X_{vw}(G^*) = j\bigr) + O\bigl(\mu_{vw}^j R/(j-1)!\bigr), \quad (5.12) \end{aligned}$$

which by summing over $j \geq 2$ yields, again using $\mu_{vw} = O(1)$ and recalling that $vw = e_1$,

$$\mathbb{E}\bigl(Z(G^*) J_{e_1}(G^*)\bigr) = \mathbb{E}\bigl(Z(G^*)\bigr)\, \mathbb{E}\bigl(J_{e_1}(G^*)\bigr) + O\bigl(\mu_{e_1}^2 R\bigr),$$

the sought estimate. □

Proof of Lemma 5.1. As said above, we use induction on $l + m$. The result is trivial if $l + m = 0$ or $l + m = 1$. If $m \geq 1$, we use Lemma 5.4; the result follows from (5.9) together with the induction hypothesis applied to $\mathbb{E}(Z(G^*))$ and the estimate $\mathbb{E}\bigl(J_{e_1}(G^*)\bigr) = O(\mu_{e_1}^2)$ from Lemma 5.2.

If $m = 0$, we study a product $\prod_{i=1}^l I_{u_i}$ of loop indicators only. We then modify the proof above, using loops instead of multiple edges in the conditionings. More precisely, we now let G_k^* be G^* conditioned on the configuration containing the k specific pairs $(u^{(2i-1)}, u^{(2i)})$, $i = 1, \dots, k$, of half-edges at u. We couple G_k^* and G_{k-1}^* as above (with obvious modifications). In this case, the switching from G_k^* to G_{k-1}^* cannot create any new loops. Hence, if $Z := \prod_{i=2}^l I_{u_i}$, we have $Z(G_k^*) \geq Z(G_{k-1}^*)$. We obtain (5.8) exactly as before, and since Lemma 5.2 still holds, this shows that Lemma 5.3 holds, now with $F_1 = \emptyset$ and the error term $O(N^{-1} \prod_{i=2}^l \mu_{u_i})$. It follows that Lemma 5.4 holds too (with J_{e_1} replaced by I_{u_1} and $F_1 = \emptyset$) by the same proof as above. This enables us to complete the induction step in the case $m = 0$ too. □

Remark 5.5. Similar arguments show that Lemmas 5.3 and 5.4, with obvious modifications, hold in this setting, where we condition on loops at u, also for $m > 0$. A variation of our proof of Lemma 5.1 would be to use this as long as $l > 0$; the result in our Lemma 5.3 then is needed only when $l = 0$, which eliminates cases (ii), (v), (vi), (vii) from the proof. On the other hand, we have to consider new cases for the loop version, so the total amount of work is about the same.

Remark 5.6. When conditioning on loops, it is possible to argue directly with conditioning on $X_u = k$, using a coupling similar to the one for G_k^* above; we thus do not need the detour with G_k^* and Lemma 3.1 used above. However, as said above, in order to treat multiple edges, the method used here seems to be much simpler. A possible alternative would be to use the methods in McKay [10] and McKay and Wormald [11, 13]; we can interpret the arguments there as showing that suitable switchings yield an approximate, but not exact, coupling when we condition on exact numbers of edges in different positions.

Remark 5.7. A small example that illustrates some of the complications when conditioning on a given number of edges between two vertices is obtained by taking three vertices $1, 2, 3$ of degree 2 each. Note that if $X_{12} = 1$,

then the multigraph must be a cycle; in particular, $X_3 = 0$. On the other hand, $X_3 = 1$ is possible for $X_{12} = 0$; this shows that it is impossible to couple the multigraphs conditioned on $X_{12} = 0$ and on $X_{12} = 1$ by moving only two edges as in the proof above. Note also that $X_3 = 0$ is possible also when $X_{12} = 2$; there is thus a surprising non-convexity.

6. The proofs are completed

Proof of Theorem 1.4. We begin by observing that the two expressions given in (1.4) and (1.5) are equivalent. Indeed, if we define Λ by (1.6), then

$$\begin{aligned}\tfrac12\sum_i \lambda_{ii} + \tfrac12\sum_{i<j}\lambda_{ij}^2 &= \tfrac12\sum_i \lambda_{ii} + \tfrac14\Bigl(\sum_{i,j}\lambda_{ij}^2 - \sum_i \lambda_{ii}^2\Bigr)\\ &= \frac12\sum_i \frac{d_i(d_i-1)}{2N} + \frac14\Bigl(\sum_{i,j}\frac{d_i(d_i-1)d_j(d_j-1)}{4N^2} - \sum_i \frac{d_i^2(d_i-1)^2}{4N^2}\Bigr)\\ &= \Lambda + \Lambda^2 - \frac{\sum_i d_i^2(d_i-1)^2}{16N^2}\end{aligned}$$

and

$$\Lambda^2 + \Lambda = \Bigl(\Lambda + \frac12\Bigr)^2 - \frac14 = \frac14\Bigl(\frac{\sum_i d_i^2}{2N}\Bigr)^2 - \frac14. \tag{6.1}$$

We note for future reference that

$$\sum_i \lambda_{ii} = \frac{1}{2N}\sum_i (d_i^2 - d_i) = \tfrac12 N^{-1}\sum_i d_i^2 - 1$$

and $\lambda - \log(1+\lambda) \ge 0$ when $\lambda \ge 0$, and thus the right-hand side of (1.4) can be estimated from above by

$$\text{right-hand side of (1.4)} \le \exp\Bigl(\frac12 - \frac{\sum_i d_i^2}{4N}\Bigr) + o(1). \tag{6.2}$$

Similarly, $\log(1+\lambda) - \lambda + \frac12\lambda^2 \ge 0$ when $\lambda \ge 0$, and thus

$$\text{right-hand side of (1.5)} \ge \exp\Bigl(-\Bigl(\frac{\sum_i d_i^2}{4N}\Bigr)^2\Bigr) + o(1). \tag{6.3}$$

In particular, since we just have shown that these two right-hand sides are the same, they tend to 0 if and only if $\sum_i d_i^2/N \to \infty$.

Next, suppose first that $\sum_i d_i^2 = O(N)$. Recall $Y(G^*) = \sum_{u\in V_n} I_u(G^*) + \sum_{e\in E_n} J_e(G^*)$ defined in (2.1). As said above, Lemma 5.1 shows that the random variables $I_u(G^*)$ and $J_e(G^*)$ are almost independent. We can compare them with truly independent variables as follows.

Let $\bar I_u$ and $\bar J_e$ be *independent* 0–1 valued random variables such that $\mathbb{P}(\bar I_u = 1) = \mathbb{P}\bigl(I_u(G^*) = 1\bigr)$ and $\mathbb{P}(\bar J_e = 1) = \mathbb{P}\bigl(J_e(G^*) = 1\bigr)$, and let

$$\overline{Y} := \sum_{u\in V_n} \bar I_u + \sum_{e\in E_n} \bar J_e.$$

Fix $k \geq 1$. We use Lemma 5.1 for all pairs (l,m) with $l+m=k$ and sum (5.4) over all such (l,m) and all distinct $u_1,\dots,u_l$ and $e_1,\dots,e_m$, multiplying by the symmetry factor $\binom{k}{l}$, and noting that the first term on the right-hand side of (5.4) can be written $\mathbb{E}\bigl(\prod_i \bar{I}_{u_i}\prod_j \bar{J}_{e_j}\bigr)$. This gives

$$\mathbb{E}\bigl(Y(G^*)^{\underline{k}}\bigr) = \mathbb{E}(\overline{Y}^{\underline{k}}) + O\Biggl(\sum\biggl(\Bigl(N^{-1}+\sum_{v\in F} d_v^{-1}\Bigr)\prod_{i=1}^{l}\mu_{u_i}\prod_{j=1}^{m}\mu_{e_j}^2\biggr)\Biggr), \quad (6.4)$$

summing over all such $l,m,(u_i)_i,(e_j)_j$ and with F depending on them as in Lemma 5.1.

Consider one term in the sum in (6.4), write as usual $e_j = v_j w_j$, and let H be the multigraph with vertices $V(H)=\{u_i\}\cup\{v_j,w_j\}$ and one loop at each u_i and two parallel edges between v_j and w_j for each $j\leq m$. Let $d_{v;H}$ be the degree of vertex v in H and note that each degree $d_{v;H}$ is even, and thus at least 2, and that F is the set of vertices with $d_{v;H}\geq 4$. We have

$$\prod_{i=1}^{l}\mu_{u_i}\prod_{j=1}^{m}\mu_{e_j}^2 = N^{-e(H)}\prod_{v\in V(H)} d_v^{d_{v;H}}, \quad (6.5)$$

where $e(H)=l+2m$ is the number of edges in H.

We group the terms in the sum in (6.4) according to the isomorphism type of H. Fix one such type $\mathcal{H}$, and let it have h vertices with degrees $a_1,\dots,a_h$ (in some order) and b edges; thus $b=\frac12\sum_{j=1}^h a_j$. The corresponding H are obtained by selecting vertices $v_1,\dots,v_h\in V_n$; these have to be distinct and it may happen that some permutations give the same H, but we ignore this, thus overcounting, and obtain from (6.5) that

$$\begin{aligned}\sum_{H\in\mathcal{H}}\Bigl(\prod_{i=1}^{l}\mu_{u_i}\prod_{j=1}^{m}\mu_{e_j}^2\Bigr) &\leq \sum_{v_1,\dots,v_h\in V_n}\Bigl(N^{-b}\prod_{j=1}^{h} d_{v_j}^{a_j}\Bigr) = N^{-b}\prod_{j=1}^{h}\Bigl(\sum_{v_j\in V_n} d_{v_j}^{a_j}\Bigr)\\ &= O\left(N^{-b+\sum_j a_j/2}\right) = O(1) \quad (6.6)\end{aligned}$$

by (2.4), since each $a_j\geq 2$.

Furthermore, let $G := \{i\in\{1,\dots,h\} : a_i\geq 4\}$. Thus, if H is obtained by choosing vertices $v_1,\dots,v_h\in V_n$, then $F=\{v_i : i\in G\}$. Consequently,

$$\begin{aligned}\sum_{H\in\mathcal{H}}\Bigl(\sum_{v\in F} d_v^{-1}\prod_{i=1}^{l}\mu_{u_i}\prod_{j=1}^{m}\mu_{e_j}^2\Bigr) &\leq \sum_{v_1,\dots,v_h\in V_n}\sum_{i\in G}\Bigl(d_{v_i}^{-1}N^{-b}\prod_{j=1}^{h} d_{v_j}^{a_j}\Bigr)\\ &= \sum_{i\in G}\Bigl(N^{-b}\prod_{j=1}^{h}\sum_{v_j\in V_n} d_{v_j}^{a_j-\delta ji}\Bigr) = O\left(N^{-b+\sum_j a_j/2-1/2}\right) = O(N^{-1/2}),\end{aligned}$$

by (2.4), since each $a_i\geq 4$ if $i\in G$ and thus $a_j-\delta_{ij}\geq 2$ for every j.

Combining this with (6.6), we see that the sum in (6.4), summing only over $H\in\mathcal{H}$, is $O(N^{-1/2})$. There is only a finite number of isomorphism

types $\mathcal{H}$ for a given k, and thus we obtain the same estimate for the full sum. Consequently, (6.4) yields

$$\mathbb{E}\big(Y(G^*)^{\underline{k}}\big) = \mathbb{E}(\overline{Y}^{\underline{k}}) + O(N^{-1/2}), \tag{6.7}$$

for every fixed k.

We use Lemma 3.2 with $\overline{Y}$ and $Y(G^*)$ (in this order). We have just verified (3.3). To verify (3.2) we take $R = 2$ (any $R < \infty$ would do by a similar argument) and find, using (5.3) and (2.2)

$$\begin{aligned}
\mathbb{E}\Big(2^{\overline{Y}}\Big) &= \prod_{u\in V_n} \mathbb{E}\, 2^{\bar{I}_u} \prod_{e\in E_n} \mathbb{E}\, 2^{\bar{J}_e} = \prod_{u\in V_n} (1 + \mathbb{E}\,\bar{I}_u) \prod_{e\in E_n} (1 + \mathbb{E}\,\bar{J}_e) \\
&\le \prod_{u\in V_n} \exp(\mathbb{E}\,\bar{I}_u) \prod_{e\in E_n} \exp(\mathbb{E}\,\bar{J}_e) = \exp\Big(\sum_{u\in V_n} \mathbb{E}\,\bar{I}_u + \sum_{e\in E_n} \mathbb{E}\,\bar{J}_e\Big) \\
&\le \exp\Big(\sum_{u\in V_n} \frac{d_u^2}{N} + \sum_{vw\in E_n} \frac{d_v^2 d_w^2}{N^2}\Big) = O(1).
\end{aligned}$$

Consequently, Lemma 3.2 applies and yields

$$\begin{aligned}
\mathbb{P}\big(G^* \text{ is simple}\big) &= \mathbb{P}(Y(G^*) = 0) = \mathbb{P}(\overline{Y} = 0) + o(1) \\
&= \prod_{u\in V_n} \mathbb{P}(\bar{I}_u = 0) \prod_{e\in E_n} \mathbb{P}(\bar{J}_e = 0) + o(1) \\
&= \exp\Big(\sum_{u\in V_n} \log \mathbb{P}(\bar{I}_u = 0) + \sum_{e\in E_n} \log \mathbb{P}(\bar{J}_e = 0)\Big) + o(1).
\end{aligned}$$

Furthermore, Lemma 4.1 yields

$$\begin{aligned}
&-\Big(\sum_{u\in V_n} \log \mathbb{P}(\bar{I}_u = 0) + \sum_{e\in E_n} \log \mathbb{P}(\bar{J}_e = 0)\Big) \\
&\qquad = \sum_{u\in V_n} \frac{d_u(d_u-1)}{4N} + \sum_{vw\in E_n} \big(-\log(1+\lambda_{vw}) + \lambda_{vw}\big) \\
&\qquad\qquad + O\left(\frac{\sum_v d_v^3}{N^2}\right) + O\left(\frac{\sum_v d_v^3 \sum_w d_w^2}{N^3}\right),
\end{aligned}$$

where the two error terms ore $O(N^{-1/2})$ by (2.4).

This verifies (1.4) and thus Theorem 1.4 in the case $\sum_i d_i^2 = O(N)$.

Next, suppose that $\sum_i d_i^2 \to \infty$. Fix a number $A > 2$. For all large N (or ν in the formulation of Theorem 1.1), $\sum_i d_i^2 > AN$, so we may assume this inequality.

Let j be an index with $d_j > 1$. We then may modify the sequence $(d_i)_1^n$ by decreasing d_j to $d_j - 1$ and adding a new element $d_{n+1} = 1$. This means that we split one of the vertices in G^* into two. Note that this splitting increases the number n of vertices, but preserves the number N of edges. We can repeat and continue splitting vertices (in arbitrary order) until all degrees $d_i \le 1$; then $\sum_i d_i^2 = \sum_i d_i = 2N$.

Let us stop this splitting the first time $\sum_i d_i^2 \le AN$ and denote the resulting sequence by $(\hat{d}_i)_1^{\hat{n}}$. Thus $\sum_i \hat{d}_i^2 \le AN$. Since we have assumed $\sum_i d_i^2 > AN$, we have performed at least one split. If the last split was at j, the sequence preceding $(\hat{d}_i)_1^{\hat{n}}$ is $(\hat{d}_i + \delta_{ij})_1^{\hat{n}-1}$, and thus

$$AN < \sum_{i=1}^{\hat{n}-1} (\hat{d}_i + \delta_{ij})^2 = \sum_{i=1}^{\hat{n}} \hat{d}_i^2 + 2\hat{d}_j + 1 - 1 \le \sum_{i=1}^{\hat{n}} \hat{d}_i^2 + 2\sqrt{AN},$$

because $\hat{d}_j^2 \le \sum_i \hat{d}_i^2 \le AN$. Consequently,

$$AN - 2\sqrt{AN} \le \sum_i \hat{d}_i^2 \le AN$$

and thus, in the limit, $\sum_i \hat{d}_i^2/N \to A$.

Let $\widehat{G}^* = G^*\big(\hat{n}, (\hat{d}_i)_1^{\hat{n}}\big)$. Since $\sum_i \hat{d}_i^2 = O(N)$, we can by the already proven part apply (1.4) to $\widehat{G}^*$ and obtain, using (6.2),

$$\mathbb{P}(\widehat{G}^* \text{ is simple}) \le \exp\left(\frac{1}{2} - \frac{\sum_i \hat{d}_i^2}{4N}\right) + o(1) = \exp\left(\frac{1}{2} - \frac{A}{4}\right) + o(1).$$

Furthermore, since $\widehat{G}^*$ is constructed from G^* by splitting vertices, $\widehat{G}^*$ is simple whenever G^* is, and thus $\mathbb{P}(G^* \text{ is simple}) \le \mathbb{P}(\widehat{G}^* \text{ is simple})$. Consequently,

$$\limsup \mathbb{P}(G^* \text{ is simple}) \le \limsup \mathbb{P}(\widehat{G}^* \text{ is simple}) \le \exp\big(-(A-2)/4\big).$$

Since A can be chosen arbitrarily large, this shows that if $\sum_i d_i^2/N \to \infty$, then $\mathbb{P}(G^* \text{ is simple}) \to 0$.

Combined with (6.2), this shows that (1.4) holds when $\sum_i d_i^2/N \to \infty$ (with both sides tending to 0).

Finally, for an arbitrary sequence of sequences $(d_i)_1^n$, we can for every subsequence find a subsequence where either $\sum_i d_i^2/N = O(1)$ or $\sum_i d_i^2/N \to \infty$, and thus (1.4) holds along the subsequence by one of the two cases treated above. It follows by the subsubsequence principle that (1.4) holds. ☐

Proof of Theorem 1.1. Part (ii) follows immediately from Theorem 1.4 and (6.2), (6.3).

If we apply this to subsequences, we see that $\liminf \mathbb{P}(G_\nu^* \text{ is simple}) > 0$ if and only if there is no subsequence along which $\sum_i (d_i^{(\nu)})^2/N_\nu \to \infty$, which proves (i). ☐

Proof of Corollary 1.5. The two expressions are equivalent by (6.1).

If $\sum_i d_i^2/N \to \infty$, then the right-hand side tends to 0, and so does the left-hand side by Theorem 1.1. Hence, by the subsubsequence principle, it remains only to show the result assuming $\sum_i d_i^2 = O(N)$. In this case we have

$$\frac{\sum_i d_i^2 (d_i - 1)^2}{16N^2} \le \frac{(\max_i d_i)^2}{N} \frac{\sum_i d_i^2}{N} = o(1)$$

and, since $\log(1+x)-x+\frac12 x^2 = O(x^3)$ for $x\ge 0$,

$$\sum_{i<j}\bigl(\log(1+\lambda_{ij})-\lambda_{ij}+\tfrac12\lambda_{ij}^2\bigr)=O\Bigl(\sum_{i,j}\lambda_{ij}^3\Bigr)$$

and

$$\sum_{i,j}\lambda_{ij}^3\le\sum_{i,j}\frac{d_i^3d_j^3}{N^3}\le\frac{(\max_i d_i)^2}{N}\Bigl(\frac{\sum_i d_i^2}{N}\Bigr)^2=o(1).$$

Hence, in this case, the formula in Corollary 1.5 follows from (1.5). □

7. Poisson approximation

As remarked in the introduction, when $\max_i d_i = o(N^{1/2})$, it is easy to prove that (1.2) implies (1.1) by the Poisson approximation method of Bollobás [2, 3]. Since this is related to the method above, but much simpler, and we find it interesting to compare the two methods, we describe this method here, thus obtaining an alternative proof of Corollary 1.5. We assume $\sum_i d_i^2 = O(N)$ throughout this section.

The main idea is to study the random variable

$$\widetilde{Y} := \sum_{u\in V_n} X_u + \sum_{e\in E_n}\binom{X_e}{2}, \tag{7.1}$$

which counts the number of loops and pairs of parallel edges (excluding loops) in G^* (we omit the argument G^* in this section). Compare this with Y defined in (2.1), and note that

$$G^* \text{ is simple} \iff \widetilde{Y}=0 \iff Y=0.$$

Theorem 7.1. *Assume that $N\to\infty$, $\sum_i d_i^2 = O(N)$ and $\max_i d_i = o(N^{1/2})$. Let $\Lambda := \frac{1}{2N}\sum_{i=1}^n\binom{d_i}{2}$ as in (1.6). Then*

$$d_{\mathrm{TV}}\bigl(\widetilde{Y},\operatorname{Po}(\Lambda+\Lambda^2)\bigr)\to 0, \tag{7.2}$$

and thus

$$\mathbb{P}\bigl(G^*(n,(d_i)_1^n) \text{ is simple}\bigr)=\mathbb{P}(\widetilde{Y}=0)=\exp\bigl(-\Lambda-\Lambda^2\bigr)+o(1). \tag{7.3}$$

If $\Lambda\to\lambda$ for some $\lambda\in[0,\infty)$, then (7.2) is equivalent to $\widetilde{Y}\xrightarrow{\mathrm{d}}\operatorname{Po}(\lambda+\lambda^2)$. (By the subsubsequence principle, it suffices to consider this case.)

Sketch of proof. We can write $\widetilde{Y}=\sum_{\alpha\in\mathcal{A}} I_\alpha+\sum_{\beta\in\mathcal{B}} J_\beta$, where $\mathcal{A}$ is the set of all pairs $\{u^{(i)},u^{(j)}\}$ of half-edges (correponding to loops), and $\mathcal{B}$ is the set of all pairs of pairs $\{\{v^{(i)},w^{(j)}\},\{v^{(k)},w^{(l)}\}\}$ of distinct half-edges (corresponding to pairs of parallel edges).

Thus, similarly to (4.1),

$$\begin{aligned}\mathbb{E}\,\widetilde{Y} &= \sum_{\alpha\in\mathcal{A}} \mathbb{E}\, I_\alpha + \sum_{\beta\in\mathcal{B}} \mathbb{E}\, J_\beta \\ &= \sum_{u\in V_n} \frac{d_u(d_u-1)}{2(2N-1)} + \frac{1}{2}\sum_{v\neq w} \frac{1}{2}\frac{d_v(d_v-1)d_w(d_w-1)}{(2N-1)(2N-3)} \\ &= \Lambda + \Lambda^2 + o(1). \end{aligned} \tag{7.4}$$

Moreover, it is easy to compute the expectation of a product of the form $\mathbb{E}\big(I_{\alpha_1}\cdots I_{\alpha_l} J_{\beta_1}\cdots J_{\beta_m}\big)$; it is just the probability that a random configuration contains all pairs occurring in $\alpha_1,\dots,\alpha_l,\beta_1,\dots,\beta_m$. If two of these pairs intersect in exactly one half-edge, the probability is 0; otherwise it is $(2N)^{-b}(1+O(1/N))$, where b is the number of different pairs. (Note that we may have, for example, $\beta_1 = \{\{v^{(1)},w^{(1)}\},\{v^{(2)},w^{(2)}\}\}$ and $\beta_2 = \{\{v^{(1)},w^{(1)}\},\{v^{(3)},w^{(3)}\}\}$, with one pair in common; thus $b \le l+2m$, but strict inequality is possible.)

We can compute factorial moments $\mathbb{E}\big(\widetilde{Y}^{\underline{k}}\big)$ by summing such expectations of products with $l+m=k$. For each term $\mathbb{E}\big(I_{\alpha_1}\cdots J_{\beta_m}\big)$, let H be the multigraph, with vertex set a subset of V_n, obtained by joining each pair occurring in $\alpha_1\dots,\beta_m$ (taking repeated pairs just once) to an edge, and then deleting all unused (i.e., isolated) vertices in V_n. It is easy to estimate the sum of these terms for a given H, and we obtain $O\big(N^{-e(H)}\prod_{v\in V(H)} d_v^{d_{v;H}}\big)$ as in (6.5). As in the proof of Theorem 1.4, we then group the terms according to the isomorphism type $\mathcal{H}$ of H. (There are more types $\mathcal{H}$ now, but that does not matter.)

Since now $\max_i d_i = o(N^{1/2})$, (2.4) is improved to

$$\sum_v d_v^k = o(N^{k/2}) \tag{7.5}$$

for every fixed $k\ge 3$, and it follows that the sum for a given $\mathcal{H}$ is $o(1)$ as soon as $\mathcal{H}$ has at least one vertex with degree ≥ 3. The only remaining case is when $\mathcal{H}$, and thus H, consists of l and m vertex-disjoint loops and double edges; in this case

$$\begin{aligned}\mathbb{E}\Big(\prod_{i=1}^{l} I_{\alpha_i}\prod_{j=1}^{m} J_{\beta_j}\Big) &= \big((2N-1)\cdots(2N-2l-4m+1)\big)^{-1} \\ &= \big(1+O(1/N)\big)\prod_{i=1}^{l}\mathbb{E}\, I_{\alpha_i}\prod_{j=1}^{m}\mathbb{E}\, J_{\beta_j}. \end{aligned} \tag{7.6}$$

Similarly, we can expand $(\mathbb{E}\,\widetilde{Y})^k = \big(\sum_{\alpha\in\mathcal{A}}\mathbb{E}\, I_\alpha + \sum_{\beta\in\mathcal{B}}\mathbb{E}\, J_\beta\big)^k$ as a sum of terms $\prod_{i=1}^{l}\mathbb{E}\, I_{\alpha_i}\prod_{j=1}^{m}\mathbb{E}\, J_{\beta_j}$ with $l+m=k$. (Now, repetitions are allowed among α_i and β_j.) If we introduce H and $\mathcal{H}$ as above, we see again that the sum of all terms with a given $\mathcal{H}$ is $o(1)$ except when $\mathcal{H}$ consists of l and

m vertex-disjoint loops and double edges. The terms occurring in this case are the same as in (7.6), and hence their sums differ by $O(1/N)$ only (since these sums are $O(1)$, see (7.4)).

Consequently, summing over all $\mathcal{H}$ and using (7.4), for every $k \geq 1$,

$$\mathbb{E}\big(\widetilde{Y}^{\underline{k}}\big) = \big(\mathbb{E}\,\widetilde{Y}\big)^k + o(1) = (\Lambda + \Lambda^2)^k + o(1).$$

If $\Lambda \to \lambda$, this shows $\widetilde{Y} \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda + \lambda^2)$ by the method of moments. In general, we obtain (7.3) and (7.2) by Lemma 3.2 and Remark 3.3. $\square$

Remark 7.2. This argument further shows that, asymptotically, the number of loops is $\mathrm{Po}(\Lambda)$ and the number of pairs of double edges is $\mathrm{Po}(\Lambda^2)$, with these numbers asymptotically independent.

In order to compare this method with the one in the preceding sections, note that $\widetilde{Y} \geq Y$ and that $\widetilde{Y} = Y$ if and only if there are no double loops or triple edges. It is easy to see that if $\sum_i d_i^2 = O(N)$ and $\max_i d_i = o(N^{1/2})$, then, using (5.2) and (7.5),

$$\begin{aligned}\mathbb{P}(\widetilde{Y} \neq Y) &\leq \sum_{u \in V_n} \mathbb{E}(X_u^{\underline{2}}) + \sum_{e \in E_n} \mathbb{E}(X_e^{\underline{3}}) \\ &= O\left(\frac{\sum_u d_u^4}{N^2}\right) + O\left(\frac{\big(\sum_v d_v^3\big)^2}{N^3}\right) = o(1), \qquad (7.7)\end{aligned}$$

so in this case the two variables are equivalent asymptotically. In particular, Theorem 7.1 is valid for Y too. It is evident that the argument to estimate factorial moments of $\widetilde{Y}$ in the proof of Theorem 7.1 is much shorter that the argument to estimate factorial moments of Y in the preceding sections. The reason for the difference is the ease with which we can compute $\mathbb{E}(I_{\alpha_1} \cdots J_{\beta_m})$ for a random configuration. Hence the proof of Theorem 7.1 is preferable in this case.

On the other hand, if $\max d_i = \Theta(N^{1/2})$, still assuming $\sum_i d_i^2 = O(N)$, there are several complications. Let us for simplicity assume that $d_1 \geq d_2 \geq \dots$, and that $d_1 \sim c_1 N^{1/2}$ with $c_1 > 0$. Then $X_1 \xrightarrow{\mathrm{d}} \mathrm{Po}(c_1^2/4)$, so $\lim \mathbb{P}(X_u > 1) > 0$ and (7.7) fails.

Moreover, cf. (7.4),

$$\mathbb{E}\,\widetilde{Y} = \tfrac{1}{2}\sum_i \lambda_{ii} + \tfrac{1}{2}\sum_{i<j} \lambda_{ij}^2 + o(1),$$

so we can write (1.4) and (1.5) as

$$\mathbb{P}\big(G^* \text{ is simple}\big) = \exp\Big(-\mathbb{E}\,\widetilde{Y} + \sum_{i<j}\big(\log(1+\lambda_{ij}) - \lambda_{ij} + \tfrac{1}{2}\lambda_{ij}^2\big)\Big) + o(1).$$

Except in the case $d_2 = o(N^{1/2})$, we cannot ignore the terms with λ_{ij} in the exponent; if, say, $d_2 \sim c_2 N^{1/2}$ with $c_2 > 0$, then $\lambda_{12} \to c_1 c_2/2 > 0$. Consequently, Theorem 1.4 shows that in this case, $\mathbb{P}(\widetilde{Y} = 0) = \mathbb{P}(G^* \text{ is simple})$

is not well approximated by $\exp(-\mathbb{E}\,\widetilde{Y})$, which shows that $\widetilde{Y}$ is *not* asymptotically Poisson distributed. (The reason is terms like $\binom{X_{12}}{2}$ in (7.1), where $X_{12} \xrightarrow{\mathrm{d}} \mathrm{Po}(c_1c_2/2)$.)

Further, we have shown in Section 6 that Y asymptotically can be regarded as the sum $\overline{Y}$ of independent indicators, but in this case $\lim \mathbb{P}(I_1 = 1) > 0$, and thus these indicators do not all have small expectations; hence Y is not asymptotically Poisson distributed.

Any attempt to show Poisson convergence of either $\widetilde{Y}$ or Y is thus doomed to fail unless $\max_i d_i = o(N^{1/2})$. It seems difficult to find the asymptotic distribution of $\widetilde{Y}$ directly; even if we could show that the moments converge, the moments grow too rapidly for the method of momemts to be applicable (at least with the Carleman criterion, see Section 4.10 in [6]). This is the reason for studying Y above; as we have seen above, the distribution is asymptotically nice, even if our proof is rather complicated.

Acknowledgement. I thank Bela Bollobás and Nick Wormald for helpful comments.

References

[1] E. A. Bender & E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A*, **24** (1978), no. 3, 296–307.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Comb.* **1** (1980), 311–316.

[3] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge Univ. Press, Cambridge, 2001.

[4] C. Cooper, The cores of random hypergraphs with a given degree sequence. *Random Structures Algorithms* **25** (2004), no. 4, 353–375.

[5] C. Cooper & A. Frieze, The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combin. Probab. Comput.* **13** (2004), no. 3, 319–337.

[6] A. Gut, *Probability: A Graduate Course*. Springer, New York, 2005.

[7] S. Janson & M. Luczak, A simple solution to the k-core problem. *Random Structures Algorithms*, to appear. `http://arxiv.org/math.CO/0508453`

[8] S. Janson, T. Łuczak & A. Ruciński, *Random Graphs*. Wiley, New York, 2000.

[9] B. D. McKay, Asymptotics for 0-1 matrices with prescribed line sums. *Enumeration and design (Waterloo, Ont., 1982)*, pp. 225–238, Academic Press, Toronto, ON, 1984.

[10] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin.* **19A** (1985), 15–25.

[11] B. D. McKay & N. C. Wormald, Uniform generation of random regular graphs of moderate degree. *J. Algorithms* **11** (1990), no. 1, 52–67. 05C80 (68R10)

[12] B. D. McKay & N. C. Wormald, Asymptotic enumeration by degree sequence of graphs of high degree. *European J. Combin.* **11** (1990), no. 6, 565–580.

[13] B. D. McKay & N. C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica* **11** (1991), no. 4, 369–382.

[14] N. C. Wormald, Some problems in the enumeration of labelled graphs. Ph. D. thesis, University of Newcastle, 1978.

[15] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B* **31** (1981), no. 2, 168–182.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se

URL: http://www.math.uu.se/~svante/